English version of "Die erste Welle der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). Stichprobendesign, Response, Gewichtung und Repräsentativität"
Bundesgesundheitsbl 2013 · 56:620–630
DOI 10.1007/s00103-012-1650-9

P. Kamtsiuris · M. Lange · R. Hoffmann · A. Schaffrath Rosario · S. Dahm · R. Kuhnert · B.M. Kurth
Department of Epidemiology and Health Monitoring, Robert Koch Institute, Berlin

The first wave of the German Health Interview and Examination Survey for Adults (DEGS1)

Sampling design, response, weighting, and representativeness

Background and purpose

The "German Health Interview and Examination Survey for Adults" (DEGS) is an integral part of the national health monitoring system that the Robert Koch Institute (RKI) carries out on behalf of the Federal Ministry of Health [1, 2].

The goal of DEGS is to regularly collect data that are representative of the whole country regarding the health situation of adults aged 18–79 years living in Germany. The data collection spectrum comprises diverse information on health status, health behaviour, living conditions, and utilisation of health care services. The information is gathered through interviews and—at certain intervals—by means of physical examinations, tests and laboratory analyses of blood and urine samples. Compared to interviews, the examination component enables more valid measurements and better prevalence estimates for diseases. A comparison of the DEGS data of different waves allows analyses of time trends. In addition, DEGS contains a panel component, i.e. participants are repeatedly invited to take part in the survey. The longitudinal data collected in this way provide insights into the causes and conditions of health changes in a person's life. The study concept is described in detail elsewhere [3, 4].

The DEGS data are included in the government's health reporting carried out at the RKI. They are used for epidemiological and public health research and made available to the scientific community by the Health Monitoring Research Data Centre at the RKI (public use files). The study results provide an important basis for health policy planning as well as evidence-based strategies of prevention and intervention measures.

The first wave of the study (DEGS1) was designed as an interview and examination survey and was conducted by the RKI from November 2008 to December 2011 [5, 6]. The data collection was carried out by two mobile examination teams under the guidance of physicians in 180 examination centres all over the country. The data collection programme was divided into two age groups (18–64 years, and older than 64). The interview part included a health and nutrition questionnaire as well as a computer-assisted medical interview and an automated assessment of currently used medications. For both age groups, the examination part consisted of laboratory analyses of blood and urine samples, an anthropometric measurement (including body height and weight) as well as blood pressure, pulse and thyroid volume measurements. For participants aged 18–64 years, an endurance bicycle ergometry was additionally conducted, whereas persons aged 65 or older completed various physical and one cognitive functional test. The articles in this issue of the *Bundesgesundheitsblatt* provide comprehensive coverage of the first results.

This overview aims to provide information for a better understanding of the individual DEGS1 result publications and to facilitate interpretation of the results. For this purpose, the sampling design, the invitation procedure, and the measures taken to ensure high participation rates in the DEGS1 are explained. Thereafter, participant numbers, reasons for non-eligibility, participation quota (response) and reasons for non-participation are presented. Since DEGS1 aims to collect data representative of the country as a whole, the participant group is compared (based on data from a short interview with non-participants) with the group of non-participants and with the German resident population (based on official statistical data). These comparisons make it possible to estimate the representativeness of the overall sample. Last but not least, the calculation of the various sample weights is described. These weights are necessary for correcting deviations of the sample from the population structure with regard to age, sex, and other features and/or to compensate for the different willingness of former GNHIES98 participants to take part in the study again.

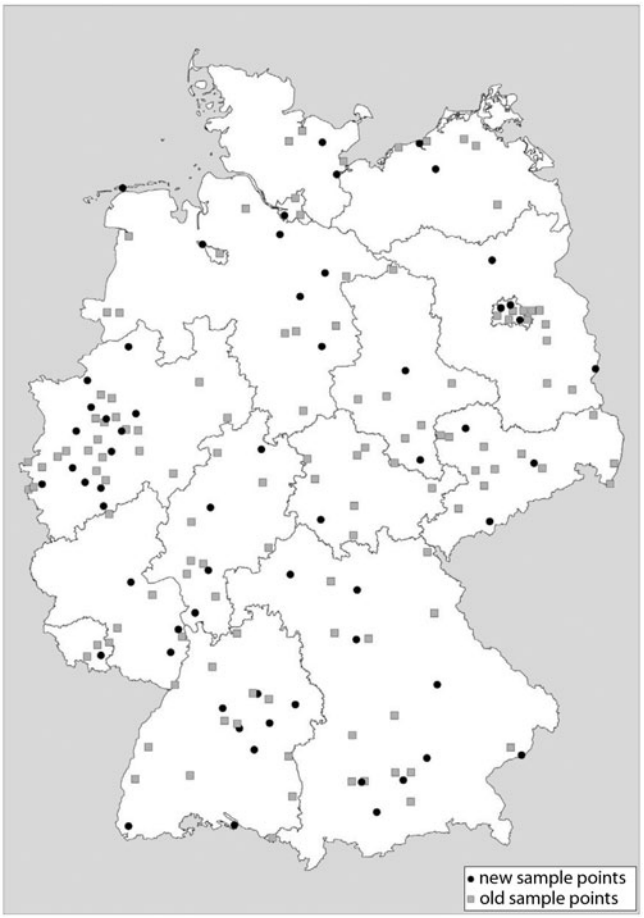

Fig. 1 ◀ DEGS1 sample points

Methods

Sampling design

The target population of DEGS1 comprised adults aged 18–79 years (basic population) who were living in Germany during the survey period and who were registered with local population registries, having indicated Germany as their main place of abode. Apart from the German population, persons of foreign nationality whose main residence was in Germany were also taken into account. People living in institutions such as residential homes were not excluded from participation. To ensure that the basic population was represented, a two-phase stratified (cluster) sample was taken in cooperation with GESIS (Leibniz Institute for the Social Sciences) in Mannheim.

First stage: selection of sample points

In the first stage of sampling, 180 sample points (i.e. study locations) were selected from the total number of federal municipalities in Germany (see **Fig. 1**). As part of this process, the 120 sample points of GNHIES98 were retained and supplemented with 60 new ones in order to reflect the current municipality structure and also to reduce design effects. This notably applies to the sample points in East Germany, since significant demographic shifts have occurred there in some places. To select the sample points both within the framework of GNHIES98 and for the 60 new points, all federal municipalities were stratified by federal state and type of municipality (10-stage BIK Classification [7]). This resulted in stratification cells. Each stratification cell was assigned a stratification weighting factor that was proportional to the overall population (of those aged 18 years and older) for all municipalities contained in the relevant stratification cell. With the help of a suitable drawing procedure [8], stratification cells were selected randomly, although with a probability proportional to the stratification weight. For each cell, the number of municipalities to be selected was determined. The municipality(ies) for each cell was(were) then drawn with probability proportional to the population. This procedure ensures that each municipality in Germany was drawn with a probability proportional to its number of inhabitants and that the number of drawn municipalities at the district, administrative region, and federal state level approximately corresponds to what was to be expected from the population at the various regional levels.

In GNHIES98, an oversampling was conducted in East Germany. This means that a disproportionate number of sample points were drawn in East Germany in order to increase the power for comparisons between east and west. In view of the increasing assimilation of the living conditions in the east and west, the new sample points were drawn without such oversampling. Because the 120 sample points of GNHIES98 were kept, this means that there is now only a slight oversampling of East Germany in the overall data set.

Very large municipalities (e.g. Berlin and Hamburg) are represented in the sample with several sample points.

Second stage: selection of target persons

In the second step, men and women aged 18–79 years were invited to participate in the survey. The sample was selected at random and stratified by age groups on the basis of the population registries of the 180 sample points. In the 120 sample points that had already been used in GNHIES98, the expected number of persons from GNHIES98 willing to re-participate was taken into account in determining the number of new persons to be drawn. As a result, a reduced number of new persons were drawn. Addresses were selected by means of a mathematical random procedure (unrestricted random selection) from the address files of local population registries. In order to compensate for the lower participation

rate expected in the larger municipalities, the number of adults to be drawn was dependent on community size. In addition, oversampling by a factor of 1.5 was performed for persons without German nationality, in order to compensate for the empirically lower participation rate and also the higher proportion of non-eligibility for this group [9]. The goal of oversampling was to ensure that the proportion of foreigners among participants corresponds to the percentage in the population.

Invitation of the participants

Over the 3-year study period, the sample points were visited according to a pre-determined sequence (see [5]). For each examination centre, the multi-stage process of participant invitation spanned about 5 weeks. The process comprised an invitation and reminder letter as well as so-called regional preparatory field visits in which invited persons who had neither accepted nor refused participation were contacted by phone and/or a home visit.

The GNHIES98 participants who had been re-invited and who declined a visit to the DEGS1 examination centre or who cancelled the agreed examination appointment were asked to participate in a questionnaire programme conducted in writing and by phone (see [5]). The paper-and-pencil questionnaires were sent to the participants by post or they were given to them as part of the regional preparatory field visits. They were reminded by means of a letter or phone call to return the questionnaires.

The GNHIES98 participants no longer living at their former municipality were not invited to an examination centre. Instead, a written invitation was sent to them in which they were asked to participate in the above-mentioned questionnaire programme. A questionnaire was already included with the invitation letter and, where applicable, with the subsequent reminder letter.

Measures for improving participation

In order to maximise the participation and utilisation rates for the study centres, different measures were used. These included, for example, ensuring that address data were as up to date as possible. This was achieved by local population registries drawing samples as near as possible to the time of invitation. Other measures taken were survey-related local public relations work [5], setting up a toll-free survey number for invited persons, regional preparatory field visits and expense allowances for the participants in the form of cash (incentives).

The participation rates were continuously observed. In order to be able to take measures to increase participation during the course of the study, the par-

Abstract · Zusammenfassung

Bundesgesundheitsbl 2013 · DOI 10.1007/s00103-012-1650-9

P. Kamtsiuris · M. Lange · R. Hoffmann · A. Schaffrath Rosario · S. Dahm · R. Kuhnert · B.M. Kurth

The first wave of the German Health Interview and Examination Survey for Adults (DEGS1). Sampling design, response, weighting, and representativeness

Abstract

The "German Health Interview and Examination Survey for Adults" (DEGS) is part of the health monitoring system of the Robert Koch Institute (RKI) and is designed as a combined cross-sectional and longitudinal survey. The first wave (DEGS1; 2008–2011) comprised interviews and physical examinations. The target population comprised 18- to 79-year-olds living in Germany. The mixed design consisted of a new sample randomly chosen from local population registries that was supplemented by participants from the "German National Health Interview and Examination Survey 1998" (GNHIES98). In total, 8,152 persons took part, among them 4,193 newly invited participants (response 42%) and 3,959 who had previously taken part in GNHIES98 (response 62%). In all, 7,238 participants visited one of the 180 local study centres, and 914 took part in the interview-only programme. A comparison of the net sample with the group of non-participants and with the resident population of Germany suggests a high representativeness regarding various attributes. To account for certain aspects of the population structure, cross-sectional, trend and longitudinal analyses were corrected by weighting factors. Furthermore, different participation probabilities of the former participants of GNHIES98 were compensated for.

Keywords

Health survey · Adults · Sample · Response · Representativeness

Die erste Welle der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). Stichprobendesign, Response, Gewichtung und Repräsentativität

Zusammenfassung

Die „Studie zur Gesundheit Erwachsener in Deutschland" (DEGS) ist Bestandteil des Gesundheitsmonitorings des Robert Koch-Instituts (RKI) und als kombinierte Quer- und Längsschnitterhebung konzipiert. Die erste Erhebungswelle (DEGS1; 2008–2011) umfasste Befragungen und Untersuchungen. Zielpopulation waren die in Deutschland lebenden Erwachsenen bis zum Alter von 79 Jahren. Das Mischdesign der Studie umfasste eine neue Einwohnermeldeamtsstichprobe, die durch Teilnehmende des Bundes-Gesundheitssurveys 1998 (BGS98) ergänzt wurde. Insgesamt nahmen 8152 Personen teil, darunter 4193 Ersteingeladene (Response 42%) und 3959 Wiedereingeladene (Response 62%). 7238 Personen besuchten eines der 180 Untersuchungszentren, 914 wurden ausschließlich telefonisch/schriftlich befragt. Der Vergleich verschiedener Merkmale zwischen der Nettostichprobe, der Gruppe der Nichtteilnehmer und der Bevölkerung Deutschlands weist auf eine hohe Repräsentativität hin. Gewichtungsfaktoren wurden berechnet, um Querschnitt-, Trend- und Längsschnittanalysen hinsichtlich einzelner Merkmale der Bevölkerungsstruktur zu korrigieren. Ferner wird bei den ehemaligen BGS98-Teilnehmenden die unterschiedliche Wiederteilnahmebereitschaft ausgeglichen.

Schlüsselwörter

Gesundheitssurvey · Erwachsene · Stichprobe · Response · Repräsentativität

Tab. 1 Overview with key figures for recruiting DEGS1 participants

		Old sample (former participants of GNHIES98)	New sample from population registries	Total
GNHIES98 participants		7,124	–	
Not eligible 1 (deceased, abroad, cannot be found) before start of study in 2008		569	–	
Unadjusted gross sample		6,555	11,008	
Not eligible 2 (deceased, gone abroad) in the course of the field work		197	1,061	
Adjusted gross sample (18–91 years)		6,358	–	
Adjusted gross sample (18–79 years)		5,927	9,947	
Non-participants		3,313	5,754	
Participants (net sample)				
Total sample (up to age 91)	Examination and interview	3,045	4,193	7,238
	Only interview	914	–	**914**
	Total	**3,959**	**4,193**	**8,152**
	Response rate	*62%*	*42%*	
Representative cross-sectional part (up to age 79)	Examination and interview	2,923	4,193	7,116
	Only interview	872	–	872
	Total	**3,795**	**4,193**	**7,988**
	Response rate	*64%*	42%	
Longitudinal study (28–91 years)	Examination and interview	3,045	–	–
	Only interview	914	–	–
	Total	**3,959**	–	–

Tab. 2 DEGS1 overall net sample, differentiated by age, sex, first and repeat participation

Age in years	18–29	30–39	40–49	50–59	60–69	70–79	18–79	80–91	18–91
Women									
Repeat participants	5	222	494	500	452	349	**2,022**	85	**2,107**
First-time participants	542	319	329	357	341	288	**2,176**	0	**2,176**
Total	547	541	823	857	793	637	**4,198**	85	**4,283**
Men									
Repeat participants	14	205	407	428	415	304	**1,773**	79	**1,852**
First-time participants	512	268	309	307	329	292	**2,017**	0	**2,017**
Total	526	473	716	735	744	596	**3,790**	79	**3,790**
Total									
Repeat participants	19	427	901	928	867	653	**3,795**	164	**3,959**
First-time participants	1,054	587	638	664	670	580	**4,193**	0	**4,193**
Total	1,073	1,014	1,539	1,592	1,537	1,233	**7,988**	164	**8,152**

ticipants were asked, for example, in case of cancellation, for the reasons for non-participation. In addition, participants were asked, as part of an additional interview at the end of the examination appointment, to provide information on their level of satisfaction with the service. As a consequence of this, appointment times were improved, the time required for examinations reduced and the level of expense allowances increased. With a view to collecting longitudinal data, it was important to involve as many GNHIES98 participants as possible. This was only possible by constantly updating the address data through researching the population registries and commercial address service providers. In addition, former GNHIES98 participants were offered participation in an interview programme, if they no longer lived at their former survey location or if they still lived there but did not want or were not able to go to their examination centre.

To ensure that the sample is representative, it is important that all groups of the population are involved in the study to a sufficient degree. This notably applies to groups of persons for whom lower participation rates are to be expected. In order to improve access for persons with limited mobility, the possibility of compensation for taxi costs and taking along an accompanying person were pointed out in the recruitment process. For people who were not able to express their consent, it was possible to be accompanied by their legal representative (that representative signing the participation consent form and, where applicable, answering interview questions on behalf of the test person, i.e. they acted as proxy). As part of the process, the examination and interview programme was adjusted to the given possibilities of a participant in accordance with a phase model. Special importance was also attached to persons with a migration background. However, for legal and ethical reasons, participants needed to have a command

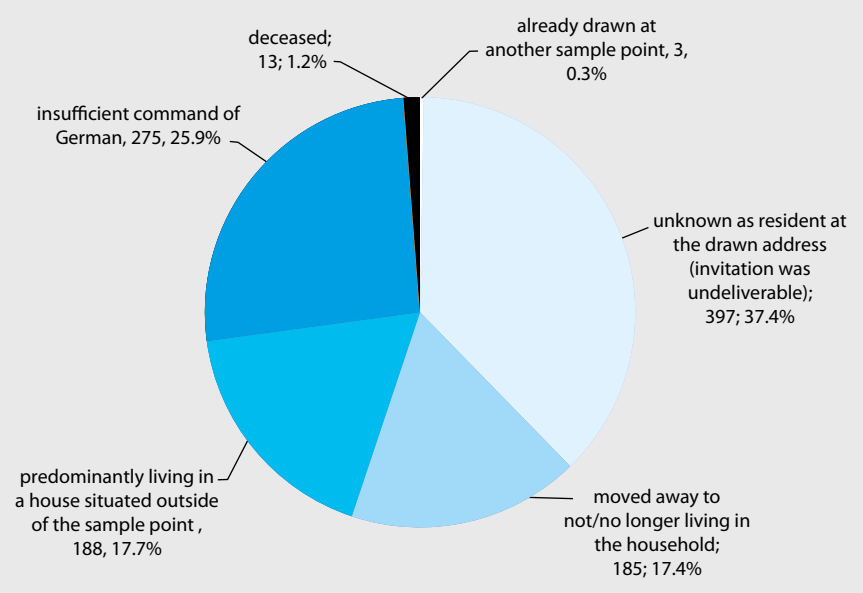

Fig. 2 ▲ Composition of the non-eligible group (n=1,061) in the newly drawn sample

of German that went beyond "broken German" to ensure that they were able to understand the verbal instructions given during the endurance bicycle ergometry and when blood was taken and also to make sure that they could express themselves in case of problems. In order to lower the language barrier, translations of the consent forms and of the health questionnaires were offered in English, Russian, Serbo-Croatian and Turkish.

Results

Invited persons

At the 180 sample points, 11,008 men and women aged 18–79 years were selected from the local population registries (see **Tab. 1**) and invited to participate in the study (adjusted gross sample). Of these participants, 1,387 did not have German nationality (including 430 from "foreign nationals oversampling").

Of the 7,124 former GNHIES98 participants, 458 had deceased before the RKI started sending out invitations, 52 had moved abroad and 59 could no longer be tracked down. Deducting these 569 persons as non-eligible resulted in an unadjusted gross sample of 6,555 former GNHIES98 participants at the beginning of the field work. Of these, 145 had refused to be contacted again during the GNHIES98 and were therefore counted as non-participants for the purpose of DEGS1. As a result, the number of GNHIES98 participants effectively to be invited was reduced to 6,410. Of these, 8 persons were not taken into account for administrative reasons, meaning that the number of participants for re-invitation amounted to 6,402 (aged 28–91 years).

Thus, a total of 17,410 men and women aged 18–91 years were invited.

Non-eligible persons in the course of the field work

Of the 11,008 newly invited persons, 1,061 (9.6%) were classified as not eligible (see **Fig. 2**). In GNHIES98 this proportion amounted to 12.3% [10]. For DEGS1, this resulted in an adjusted gross sample of 9,947 newly invited adults. Persons who were unknown as residents at the drawn address (37.4%) or who had moved away (17.4%) accounted for the largest shares of the non-eligible group, totalling more than half. Another 17.7% predominantly lived in houses situated outside the sample point. In addition, 25.9% of the non-eligible group concerned persons who, due to an insufficient command of German, were not eligible to participate in the study. Situations where persons had already been drawn at another sample point or had deceased in the meantime were rare (0.3 and 1.2%, respectively).

For re-invited persons (unadjusted gross sample n=6,555), subjects were defined as not eligible—in addition to the"not eligible" status established already before the start of the study—if it only became known in the course of the invitation that they had moved abroad (n=7; 0.1%) or had deceased (n=190; 2.9%). This resulted in a gross sample of 6,358 re-invited persons.

If the non-eligible group resulting from the preliminary research is also taken into account, the proportion of deceased persons in the GNHIES98 overall sample amounts to 9.1%, and at 0.7–0.8% corresponds to the mortality rate per year shown in the official statistics [11]. The RKI intends to research the cause of death for the total of 648 deceased GNHIES98 participants and to conduct mortality analyses. The feasibility of research into the cause of death has already been ascertained in a pilot project [12].

Participants

A total of 8,152 persons aged 18–91 years participated in DEGS1 (see **Tab. 2**). Of these individuals, 315 did not have German citizenship, 4,193 had been invited for the first time and 3,959 had already participated in GNHIES98.

Of the 8,152 participants, 7,238 visited one of the 180 examination centres and completed both the examination and the interview programme there (see **Tab. 1**). This results in an average number of 40.2 study participants per sample point. Whereas—as shown in **Tab. 1**—all of the 4,193 newly invited participants visited one of the examination centres, out of those who were re-invited, 3,045 participated in the study at the examination centre. Another 914 re-invited participants only completed the interview programme (no visit to an examination centre)—419 of them no longer lived at their former survey location. The number of cases for cross-sectional and longitudinal analyses resulting from this sample is shown in **Tab. 1**.

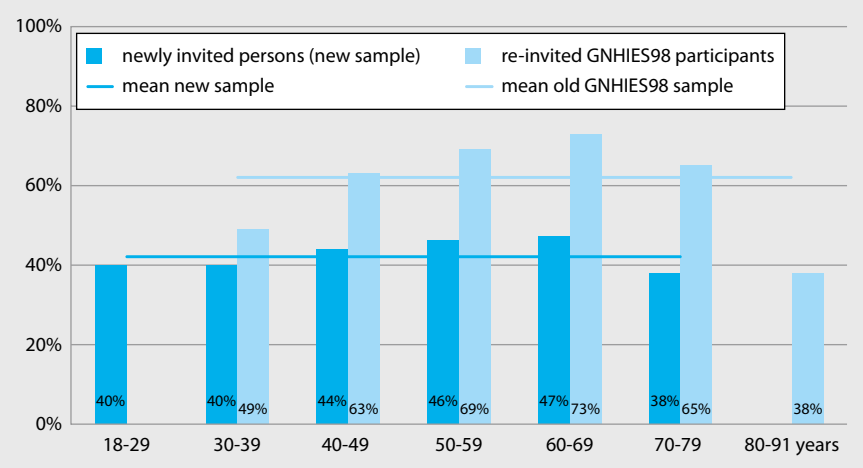

Fig. 3 ▲ Response by age for newly invited and re-invited persons

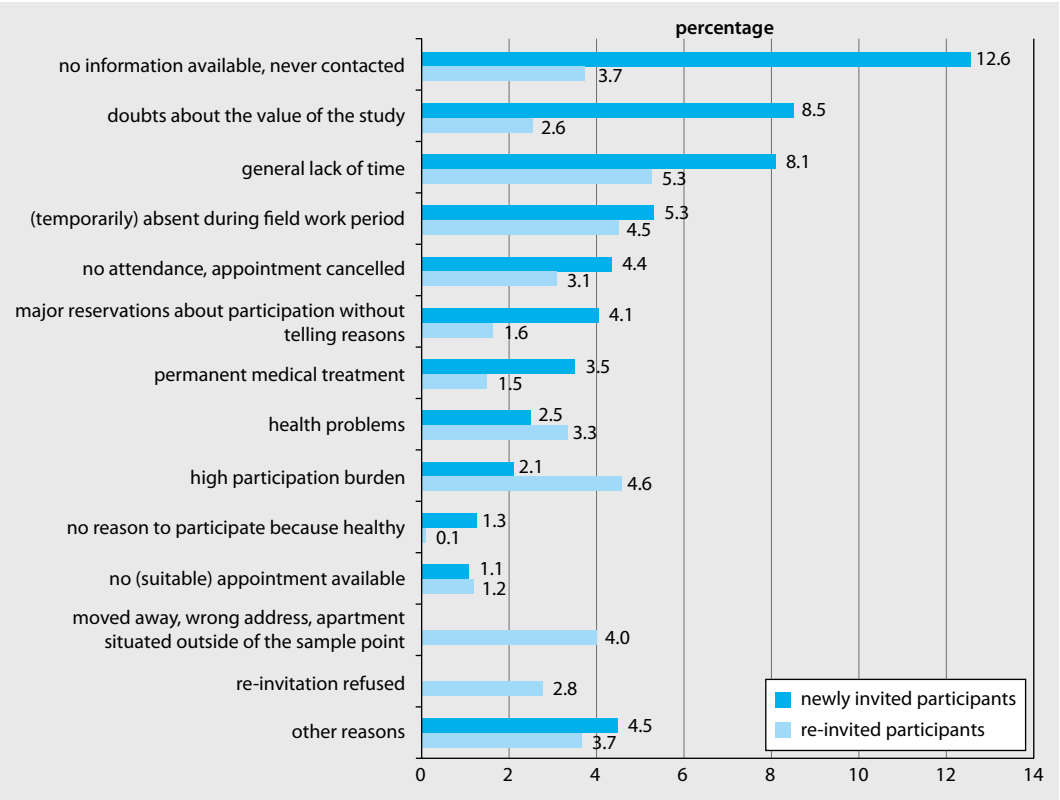

Fig. 4 ◀ Reasons for non-participation in the examination part for newly invited persons (adjusted gross sample, n=9,947) and re-invited persons (adjusted gross sample minus persons invited directly to the interview programme, n=5,247)

Response rate

In DEGS1, a response rate of 62% was reached for former GNHIES98 participants—excluding those aged 80–91, the response rate is 64%. In the GNHIES98, the response rate was 61% [10]. Among newly invited persons, the response rate was 42%. Compared to other national surveys conducted in Europe, this response rate is average [13].

When the willingness to participate is analysed by age group, striking differences are found only in the sample of re-invited persons (see **Fig. 3**). Thus, participation was above average for the age group of 50- to 69-year-olds, whereas the participation rates for the younger age group of 30–39 years as well as for persons aged 80 or older was below average. However, the reasons for the lower willingness to participate are different for the two age groups. Younger persons predominantly have no time or interest, whereas for the elderly, the effort involved leads to non-participation. The different non-response rates in the various age groups were corrected by the sample weights (see below).

Non-response analysis and representativeness

Fig. 4 gives an overview of the reasons for non-participation in the exami-

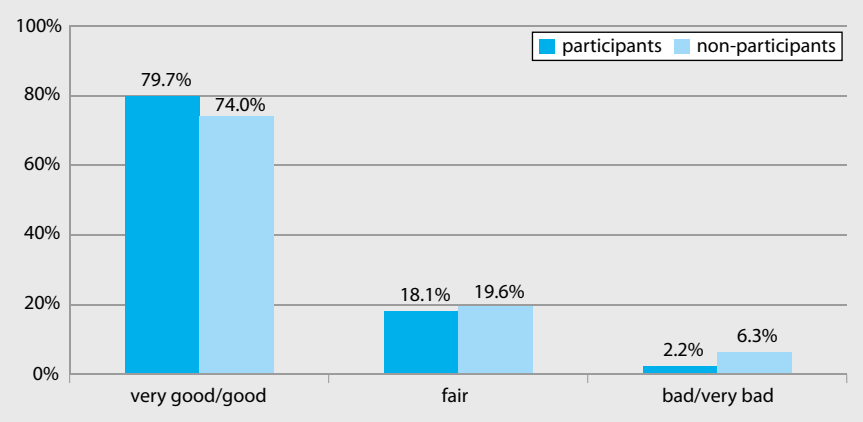

Fig. 5 ▲ Self-perceived health for newly invited participants and non-participants of the DEGS1 (both weighted: $n_{participants}$ =3,909, $n_{non\text{-}participants}$ =2,342)

nation centre. The percentages given apply to the adjusted gross sample of persons (participants and non-participants) who were invited to an examination centre (newly invited persons n=9,947, re-invited persons n=5,274). In the re-invited group, persons no longer living in their former sample point, who for that reason were directly invited to participate in the interview programme, were not included here.

The evaluation shows that lack of time was a significant reason for non-participation. This applies both to persons who did not have time temporarily and also to persons who stated that they were generally too busy to participate. Whereas the proportion of persons who were temporarily short of time was similar for those re-invited (4.5%) and those invited for the first time (5.3%), a significantly higher proportion of newly invited persons (8.1%) compared to those re-invited (5.3%) stated that they did not have time generally. It was clearly more common for those invited for the first time (8.5%) to have doubts about the value of the study. For those re-invited, who had already experienced participation, only 2.6% had doubts about the value of the study (gratifyingly, only 0.2% reported explicitly bad experiences with GNHIES98 as a reason for non-participation). A similar difference is evident for persons who appear to have had major reservations about participation in the study and who refused participation without giving other reasons (first-time invitees 4.1%; repeat invitees 1.6%). Health problems (acute or chronic illness, current hospitalisation or stay at a health resort) were again frequently cited reasons—both by newly invited persons (2.5%) and by re-invited persons (3.3%). Many people also commented that they were receiving medical treatment on a permanent basis (newly invited persons 3.5%; re-invited persons 1.5%). One aspect of the organisation of the study that turned out to be problematic was the fact that a relatively large number of persons cancelled or failed to attend their appointments. The proportion of cancelled or missed appointments (excluding persons whose appointment was rescheduled) was, at 4.4%, once again slightly higher among those newly invited than the 3.1% for re-invited persons. Despite the waiting list, it is difficult to give appointments to another person if they are cancelled at very short notice. Last but not least, for 12.6% of newly invited persons and 3.7% of those re-invited, no information was available on the reasons for non-participation, because it was never possible to contact them in person.

An important goal of DEGS1 was to achieve a sample that was representative of the resident population of 18- to 79-year-olds in Germany. In practice, the response rate has established itself as a measure for representativeness and hence for the quality of a survey. However, the bias of a population survey and hence the quality of the collected data also depend on the effective difference in characteristics between participants and non-participants (non-responders). Non-responders only pose a central problem if they vary systematically from responders. Bias in estimates cannot be excluded, even if the response rate is high [14, 15]. Rather, bias can be exacerbated, if "more of the same" is recorded. Various methodological studies have confirmed this problem [16]. In order to evaluate the representativeness of DEGS1, both indicators, i.e. the response rate and the difference between responders and non-responders, were therefore analysed. Since information about non-responders is only partially available, a comparison was additionally made between participants and the target population, with information on the latter taken from official statistics. The analysis also served the purpose of identifying characteristics that may need to be taken into consideration when calculating the sample weights and which are thus included in the statistical adjustment of the estimators (see below).

In DEGS1, a short questionnaire for non-participants was used that asks questions about both socio-demographic and health-related characteristics (e.g. body height and weight, self-perceived health, smoking status). Among the newly invited persons, 42% of non-responders answered this questionnaire. This percentage corresponded to just under a quarter of the adjusted gross sample, meaning that basic information is available from just under 70% of the gross sample. For the possible bias in survey results, a comparison of health-related characteristics between participants and non-participants is particularly important. Regarding smoking behaviour, for example, hardly any differences between responders and non-responders are detectable. Nor are there any significant differences (data not shown) for obesity and overweight (based on self-reports of height and weight). **Fig. 5** shows that even in participants' self-perceived health, there are hardly any differences.

A comparison of the overall net sample with the resident population of Germany (target population) by means of indicators from official statistics also suggests high representativeness. For the socio-demographic characteristics, the analyses show typical differences to

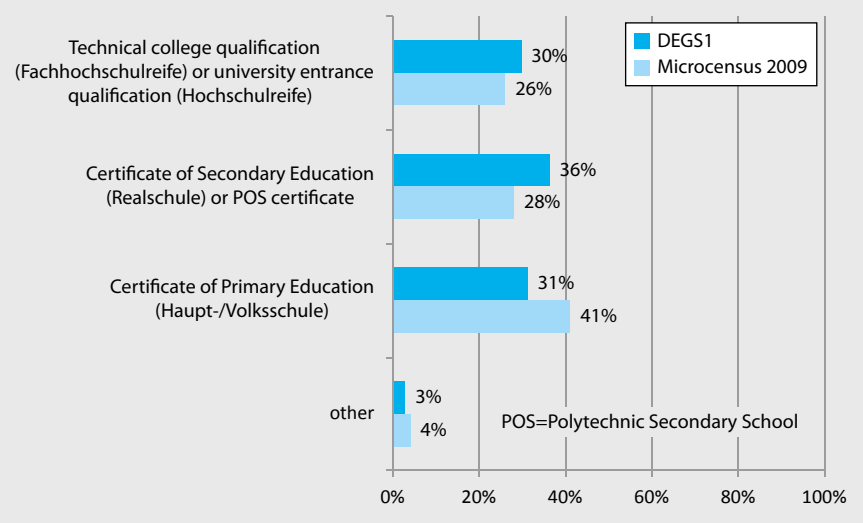

Fig. 6 ▲ Comparison of the highest school leaving certificates between the DEGS1 participants (unweighted: n=7,257) and the population

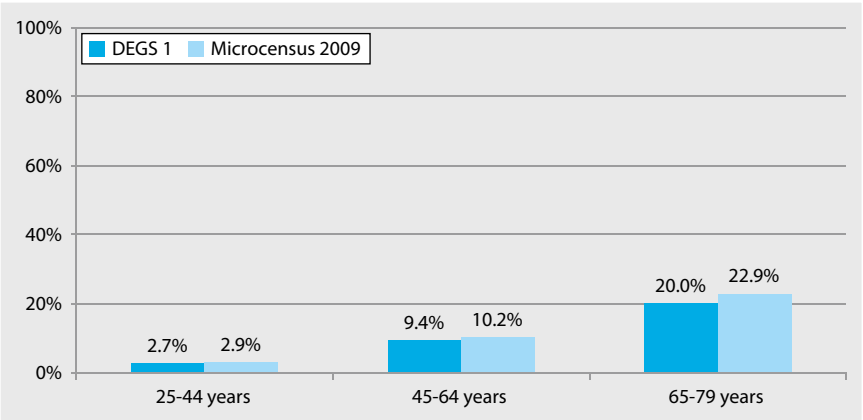

Fig. 7 ▲ Comparison of officially recognised severe disability between the DEGS1 participants (weighted: n=6,974) and the population

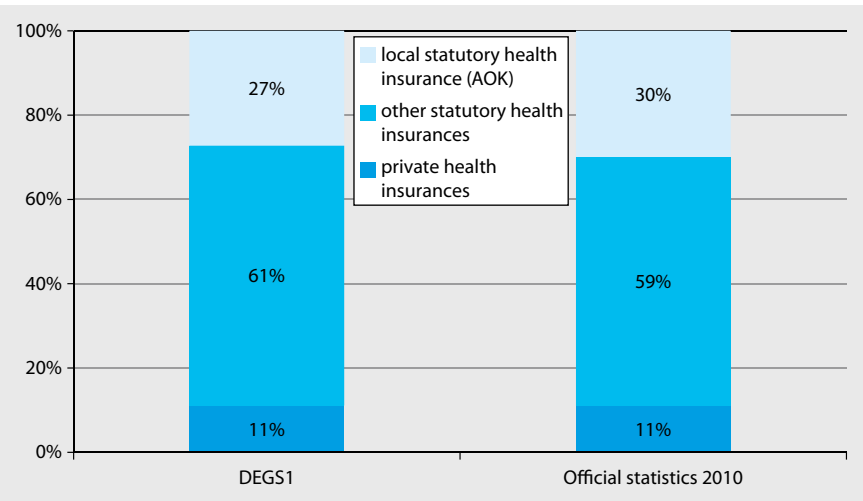

Fig. 8 ▲ Comparison of health insurance between the DEGS1 participants (weighted: n=7,757) and the population

other studies [17, 18, 19]. For example, as **Fig. 6** shows, for the characteristic "Highest School Qualification" in the DEGS1 sample, the number of people with a (vocational) university entrance qualification is almost the same as that of the general population (26% vs. 30%), whereas persons with a certificate of primary education (Haupt-/Volksschule) are less common than in the general population (31% vs. 41%) [20]. However, these differences are corrected by including educational status in the calculation of the sample weights (see below).

Two additional examples of health and/or health-care-related indicators may serve to clarify the degree to which the above-mentioned goal is achieved. **Fig. 7** shows that the proportion of the DEGS1 participants with an officially recognised severe disability roughly corresponds to the proportion that the official statistics show for the various age groups [20]. A second example concerns the proportion of persons who have statutory health insurance (Gesetzliche Krankenversicherung, GKV). Among the survey participants, this proportion is about 90% and differs only slightly from the actual proportion in the resident population (see **Fig. 8,** [21, 22]).

Sample weights

In order to be able to make representative statements about the adult population aged 18–79 years in Germany and to perform trend analyses by comparing the DEGS1 data with the GNHIES98 data, sample weights were calculated. This was done in a two-stage process. In the first step, the design weights were calculated separately for the two samples (newly invited persons and re-participants of GNHIES98). In the second step, the overall sample was adjusted to the population figures of the Federal Statistical Office for the year 2010 (as of 31 December 2010) through post-stratification. Overall, four types of sample weights are available for the analysis of the interview data (n=7,988) and the examination data (n=7,238). These can be used for cross-sectional and trend analyses. In addition, sample weights for longitudinal analyses are available. The calculation of the sam-

Tab. 3 Statistical characteristics of the sample weights in the cross-sectional examination part of the DEGS1 (n=7,116)

		Min.	1. Quartile	Median	Mean value	3. Quartile	Max.	Effectiveness [%]	<=0.3 [%]	>=3 [%]	N
Design weights First-time participants	West	0.27	0.69	0.90	1.02	1.22	3.51	82.2	0.6	0.9	3,011
	East	0.24	0.43	0.53	0.62	0.71	3.51	77.1	4.6	0.3	1,026
	Berlin	0.24	0.51	0.90	1.07	1.31	3.51	70.0	2.6	3.8	156
	Total	0.26	0.64	0.90	1.00	1.22	3.79	78.2	0.9	1.1	4,193
Design weights Repeat participants		0.21	0.53	0.87	1.00	1.22	4.91	68.3	2.8	2.2	2,923
Design total		0.24	0.59	0.87	1.00	1.22	5.43	72.4	1.9	1.7	7,116
Post-stratification weights		0.11	0.47	0.78	1.00	1.28	4.62	61.6	11.0	3.3	7,116

ple weights for conducting cross-sectional analyses is described below.

The design weights for the newly invited sample compensate for the stratification and clustering involved in the study design and also make up for non-response. These weighting factors are inversely proportional to the selection probability of persons, which is given by the product of the selection probability of the municipality with the selection probability of the person within the municipality. The selection probability for the municipality (i.e. the sample points) was estimated separately for the three regions West Germany, East Germany and Berlin and was calculated by dividing the size of the population of 18- to 79-year-olds in the sample point by the total population size of the relevant region. The selection probability of participants within a sample point is given by the number of net participants per sex and age group divided by the total number of persons living in the relevant sex and age group within the municipality. With the exception of the youngest group, 10-year intervals (18–29, 30–39, …, 70–79) were taken as a basis.

The design weight for re-invited persons is given by the reciprocal value of their participation probability in DEGS1. This participation probability is the product of the probability of having participated in GNHIES98 (design weight in GNHIES98) and the probability to participate again in DEGS1. Whereas the design weight was already calculated for GNHIES98, the re-participation probability was estimated by a logistic regression model. In this model, the re-participation was the target variable and various GNHIES98 variables were used as independent variables.

The design weights of two samples were standardised to the combined sample size of both samples, and then adjusted to the population numbers (post-stratification). In the post-stratification, the design weights of the combined sample were adjusted in an iterative procedure so that the estimated marginal distributions corresponded to the relevant population structure in the year 2010. The adjustment was made with regard to the population number of the federal states, BIK classification classes, age groups, sexes, nationalities and educational levels (the latter based on the micro-census data). The post-stratification enhances the representativeness of the estimates and partially compensates for non-response. In addition, the goal is to achieve consistent estimates for sub-populations such as regions or age groups.

The final weight was the result of the multiplication of the design weight with the post-stratification weight. **Tab. 3** shows the statistical characteristics of the sample weights, for the study population who participated in both the interview and the examination programme.

For longitudinal analyses, the focus was on transfer and/or projection of individual health changes between the two measurement times 1998 and 2008–2011. Since, in contrast to cross-sectional analyses, it is not the current status of the population in the years 2008–2011, but rather the change in status on an individual basis that is the focus, the population size in 1998 provides the basis for calculating the weighting factor. The longitudinal weight was also calculated through the multiplication of two factors: on the one hand, the design weight of GNHIES98 (inverse probability of having participated in GNHIES98) is taken into account; on the other hand, a re-participation weight is calculated to compensate for the different re-participation probabilities. However, in contrast to the cross-sectional weight, the former GNHIES98 participants who died from 1998 on were not excluded from the sample, because their death may be associated with health variables and thus constitutes a relevant category of health changes (cf. [23]). The deceased were therefore assigned a weighting factor of 1 and could thus be included in analyses. The re-participation probability was therefore estimated via a logistic regression model as described above, but for a reduced sample excluding the deceased. The re-participation weight was again given by the reciprocal value of the estimated re-participation probability. Unlike in the model for the cross-sectional weights, subjects 80 years old or older or persons who have moved abroad were included in the model for estimating the re-participation probabilities.

Conclusion

As part of the national health monitoring system, DEGS aims to collect representative data, for health reporting, politics and science, on the population living in Germany. This overview has discussed the sampling design, response rate and representativeness of DEGS1. The analysis showed that on the basis of the DEGS1 data, representative statements can be made on the health status, the

health behaviour and the medical care of people living in Germany. Reaching this goal required different complex strategies and measures within DEGS1.

The fact that it was also possible to achieve the DEGS1 goal with a combination of a newly drawn and an existing sample indicates that such a mixed design constitutes a suitable option for drawing samples, which enables simultaneous collection of cross-sectional and longitudinal data representative of the population. Distortions in the composition of the study population resulting from the study design can be compensated for by means of weighting.

While a high number of participants is not in itself a distinguishing feature of the high quality of the sample, a high participation rate is nevertheless important in order to increase the power of subgroup analyses.

Access via the local population registries worked very well in DEGS1. The share of quality-neutral losses due to invalid addresses was smaller for the newly drawn sample than it had been for GNHIES98. In part, this is certainly a consequence of reducing the time period between drawing the sample and the invitation being sent out. In the recruitment of participants, it was important to keep the number of invited persons who could not be contacted as low as possible to ensure that each person got a chance to participate. In the course of DEGS1, repeated efforts were made to ensure that this was the case. These included, for example, local public relations work and getting in touch with the invited persons as part of field preparations. An important aspect of this was making sure that a low-threshold interview programme was offered for all persons who were unable or did not want to visit an examination centre. This notably applies to former GNHIES98 participants who had moved away from the initial survey locations. In order to be able to assess the quality of the sample, it is also essential to collect information in the process of participant recruitment that allows an estimation of the composition of the group of non-participants (and the reasons for non-participation).

Overall, these insights permit conclusions on the preparation of future surveys. The RKI can develop data collection strategies and quality recording instruments that can be used additionally to ensure that quality monitoring of the sample and of non-response rates can be conducted during the field work phase and that adjustments can be made regarding the recruitment methods and target groups.

Corresponding address

P. Kamtsiuris
Department of Epidemiology and Health Monitoring, Robert Koch Institute
General-Pape-Str. 62–66, 12101 Berlin
Germany
KamtsiurisP@rki.de

Acknowledgements. The study was financed with resources from the Robert Koch Institute and the Federal Ministry of Health.

Conflict of interest. On behalf of all authors, the corresponding author states that there are no conflicts of interest.

References

1. Kurth BM (2012) Das RKI-Gesundheitsmonitoring – was es enthält und wie es genutzt werden kann. Public Health Forum 20(76):4.e1–4.e3
2. Kurth BM, Lange C, Kamtsiuris P, Hölling H (2009) Health monitoring at the Robert Koch-Institute. Status and perspectives. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 52:557–570
3. Gößwald A, Lange M, Kamtsiuris P, Kurth BM (2012) DEGS: German Health Interview and Examination Survey for Adults. A nationwide cross-sectional and longitudinal study within the framework of health monitoring conducted by the Robert Koch-Institute. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 55:775–780
4. Scheidt-Nave C, Kamtsiuris P, Gößwald A et al (2012) German Health Interview and Examination Survey for Adults (DEGS) –design, objectives and implementation of the first data collection wave. BMC Public Health 12:730
5. Gößwald A, Lange M, Dölle R, Hölling H (2013) The first wave of the German Health Interview and Examination Survey for Adults (DEGS1). Participant recruitment, fieldwork, and quality management. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 56:611–619
6. Robert Koch-Institut (Hrsg) (2009) DEGS: Studie zur Gesundheit Erwachsener in Deutschland – Projektbeschreibung. Beiträge zur Gesundheitsberichterstattung des Bundes. RKI, Berlin
7. Aschpurwis + Behrens GmbH (Hrsg) (2010) BIK-Regionen. Ballungsräume, Stadtregionen, Mittel-/Unterzentrengebiete. Methodenbeschreibung zur Aktualisierung 2010 Hamburg
8. Cox L (1987) A constructive procedure for unbiased controlled rounding. J Amer Statistical Assoc 82:520–524
9. Schenk L, Ellert U, Neuhauser H (2007) Cildren and adolescents in Germany with a migration background. Methodical aspects in the German Health Interview and Examination Survey for Children and Adolescents (KiGGS). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 50:590–599
10. Thefeld W, Stolzenberg H, Bellach B (1999) Bundes-Gesundheitssurvey: Response, Zusammensetzung der Teilnehmer und Non-Responder-Analyse. Gesundheitswesen 61:S57–S61
11. Robert Koch-Institut (Hrsg) (2011) Sterblichkeit, Todesursachen und regionale Unterschiede. Gesundheitsberichterstattung des Bundes, Heft 52. Berlin
12. Wolf I, Knopf H, Scheidt-Nave C, Kurth BM (2012) Possibilities and limitations of retrospective research on cause of death within the framework of a nationwide epidemiological study. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 55:431–435
13. Tolonen H, Koponen P, Aromaa A et al (2008) Review of health examination surveys in Europe. Publications of the National Public Health Institute 18/2008, Helsinki
14. Särndal C-E, Lundström S (2005) Estimation in surveys with nonresponse. Wiley
15. Schouten B, Cobben F (2007) R-indexes for the comparison of different fieldword strategies and data collection modes. Discussion paper 07002. In: Statistics Netherlands (Hrsg), Voorburg/Heerlen
16. Bethlehem J (2010) New developments in survey data collection methodology for official statistics. Discussion paper 10010. In: Statistics Netherlands (Hrsg), Hague/Heerlen
17. Koch A (1997) Teilnahmeverhalten beim ALBUS 1994. Köln Z Soziol Sozialpsychol 49:98–122
18. Neller K (2005) Kooperation und Verweigerung: Eine Non-Response-Studie. ZUMA-Nachrichten 57:9–36
19. Schärpler J-P (2000) Was kann man am Beispiel des SOEP bezüglich Nonresponse lernen? ZUMA-Nachrichten 46:117–150
20. Statistisches Bundesamt (2010) Leben in Deutschland – Ergebnisse des Mikrozensus. Eigenverlag, Wiesbaden
21. Bundesministerium für Gesundheit (2011) Gesetzliche Krankenversicherung. Mitglieder, mitversicherte Angehörige und Krankenstand. Jahresdurchschnitt 2010 (Ergebnisse der GKV-Statistik KM1/13). Stand. 1. April 2011
22. Verband der Privaten Krankenversicherung (2011) Zahlenbericht der Privaten Krankenversicherung 2010/2011. Köln, Berlin
23. Chaix B, Evans D, Merlo J, Suzuki E (2012) Commentary: weighing up the dead and missing: reflections on inverse-probability weighting and principal stratification to address truncation by death. Epidemiology 23:129–131 (discussion 132–127)